The Impact of Reference Linking on the Creation and Use of References/Citations

CENDI/FLICC Workshop
Library of Congress
June 21, 2000
Deb Bendig, OCLC
dbendig@oclc.org

Outline

Data considerations

Presentation of links

Appropriate copy issue

Centralized vs.. local linking

Other considerations - one slide only!

New partnerships

Summary

What is a citation?

- **A reference in an article**
- **An item in a bibliography**
- **A record in a citation or A&I database**

Interesting citation data

- **Author**
- **Title**
- **Journal/ISSN**
- **Year**
- **Volume**
- **Issue**
- **1st page number**
- **Unique ID**

Recommendation

- **When providing article data:**
  - Identify each citation
  - Put citation data elements into distinct fields
  - Include any unique article ID you have, even if not to be displayed
  - Use standards where possible
  - Include any unique article ID in the citation!
  - Include a unique article ID for the article!

Digression

The best article identifier is

- **associated with the article from the beginning of the publication process,**
- **displayed with the article, and**
- **carried in all references to the article.**

Make unique article ids ubiquitous!

Where will the links appear?

- **HTML articles**
- **Within PDF articles?**
- **HTML versions of PDF articles**
- **PS/TeX/etc. articles?**

Presentation of links

- **Should links be printable?**
  - Anything displayed in a browser prints
  - If link is durable, consider displaying
  - And, links may be different for different people, depending on rights
- **Should links be branded?**
  - May be required by full-text source
  - Good for unambiguous id of source

Where do the links come from?

Publisher adds them

- **Manual process $$$$**
- **Automated process to id citations, extract data, batch match against...existing full-text systems?**
- **Link maintenance?**
- **Record updated when new sources available?**
- **Assumes all users of data have access via specified link(s)**

Where do the links come from?

“System” adds them

- **Automated process, based on :**
  - Available citation data
  - Available matching articles
  - User rights information
- **On-the-fly or continuously updated**

Automated article linking challenges

- **Granularity**
- **Will the real article title stand up?**
- **Special characters**
- **Versions/editions/manifestations**
  - Reissued article
  - Modified article
  - Different ISSN for electronic version of journal
- **Remedy: TOCs?**

For now, assume a perfect match...

Appropriate copy issue

- **Same articles available from multiple sources**
- **Library or user has paid for access from one or more source**
- **How does a system know the best full-text links to display to a user?**

Library defines their best link

- **E-holdings consist of**
  - Multiple sources
  - Possibly some overlap
  - Possibly different source for different time periods
  - Possibly multiple formats, with preferred order
- **Can change frequently**
- **Therefore, library holdings profile**

Sources

- **Multiple sources of articles**
  - CrossRef, NIH/NCBI initiatives
  - Direct from publisher
  - Aggregator services
- **Overlapping coverage among sources**
- **Multiple formats: ASCII, HTML, PDF**
- **Varying linking mechanisms**
- **Agreements may control availability via system**

Journal Sources Profile

For each available source, track:

- **Journal Title**
- **ISSN**
- **Coverage**
- **Format(s)?**
- **Linking mechanism/requirements (query, DOI, unique ID)**
- **Authentication requirements?**

And/or, have a database of all available articles, with links

Given a link, what does the

- **An article (or TOC)**
  - "No dead links!"
  - Requires knowledge of article's existence, e.g, database of articles
- **The possibility of an article**
  - Service uses profile of source holdings and assumes all articles are available

Centralized vs. local linking

- **Centralized:**
  - System provided for libraries
  - Each system must know library profile
  - Links or associated services limited to system agreements/capabilities
- **Local (e.g., ExLibris SFX, consortia):**
  - Centralized holdings profile?'
  - Possibility of bigger pool of sources, services

Other linking considerations

- **Centralized library holdings profiling**
- **Authentication (especially remote user)**
- **Guaranteed archival access**
- **Per-article purchase**
- **Linking standards**
- **Source of article-use statistics for library**

New partnerships

Libraries ask online services for:

- Full-text coverage matching their access rights
- Currency: how quickly new full text integrated?
- Quality of linking
  - %-age dead links
  - %-age incorrectly linked articles
- Persistent URLs
- Centralized holdings profile
- One-stop shopping, with title-level control

New partnerships

Online services ask publishers for:

- Timely availability of online full text
- Data to support accurate online linking
  - Standard metadata
  - Unique article identifiers added in publishing process
  - Persistent URLs for full text
- Remote authentication solutions
- Cost-effective access to full text

New partnerships

Publishers ask for:

- Accurate rights management
- Access to more subscribers
- Continuation of revenue stream
- Tools to support data production
- Remote authentication solutions

Summary

- **Quality of linking becomes a significant system discriminator**
- **There will be no one source for all full text**
- **Standards will provide the best environment for reference linking**
- **Opportunities for new centralized services**

Questions?